Enabling Metabolic Pathways Analysis through Multi-Omics Integration
Lindsey Rappaport[1], Bishoy Kamel[2], Igor Grigoriev[2]
[1]Palomar College, [2]Lawrence Berkeley National Laboratory

WORKFORCE DEVELOPMENT & EDUCATION

BERKELEY LAB

Abstract: A metabolic pathway is a connected series of chemical reactions taking place in a live cell. These pathways are essential to maintain homeostasis within organisms. Data visualization is a critical part of understanding metabolic pathways. While most efforts have been geared to development of tools that analyze specific types of data such as transcriptomics, metabolomics, or proteomics, tools that are capable of integrating such heterogeneous datasets into meaningful representations that link their abundances to infer activation or repression of pathways under different conditions are scarce and still an active area of research.

The goal of this project was to develop a web-based platform to create and visualize metabolic pathway maps of important biological processes in fungi and algae overlaid with multi-omics data in a way that is organized and easy to comprehend/replicate. Pathway maps, obtained from public databases and literature, were drawn in Inkscape, a vector graphics editing software, to represent pathways (see Fig 1). Web application HTML (HyperText Markup Language) and pathway SVG (Scalable Vector Graphics) code were edited using Visual Studio Code, linking pathway image coordinates to specific enzymes and metabolites defined in JSON (JavaScript Object Notation) data structures (see Fig 2). Test data was generated as well as an optimized, curated database linking metabolites and enzymes to maps.

The resulting database of completed pathways (containing pathway names and enzymes/metabolites present in each pathway) amounted to over 30 metabolic pathways and allowed multi-omics data to be projected onto the maps. The pathway examined spanned lipid, carbohydrate, amino acid, nucleotide, secondary metabolism, and chlorophyll biosynthesis. While pathways varied in complexity and size, a stepwise streamlined drawing process was developed to handle various common metabolic pathways layouts such as cycles and linear pathways.

The diversity of pathways covered in this project combined with the embedded web interface to lay out multi-omics data will jumpstart the multi-omics capabilities planned for integration into the MycoCosm and PhycoCosm portals at the U.S. Department of Energy Joint Genome Institute (JGI).

Project Goal: Develop a web-based platform to create and visualize metabolic pathway maps integrated with metabolomics and transcriptomics data in an organized, comprehensible, and replicable way.

Figure 1: Examples of pathways drawn in Inkscape. Enzymes shown in orange, cofactors shown in blue, metabolites shown in grey.
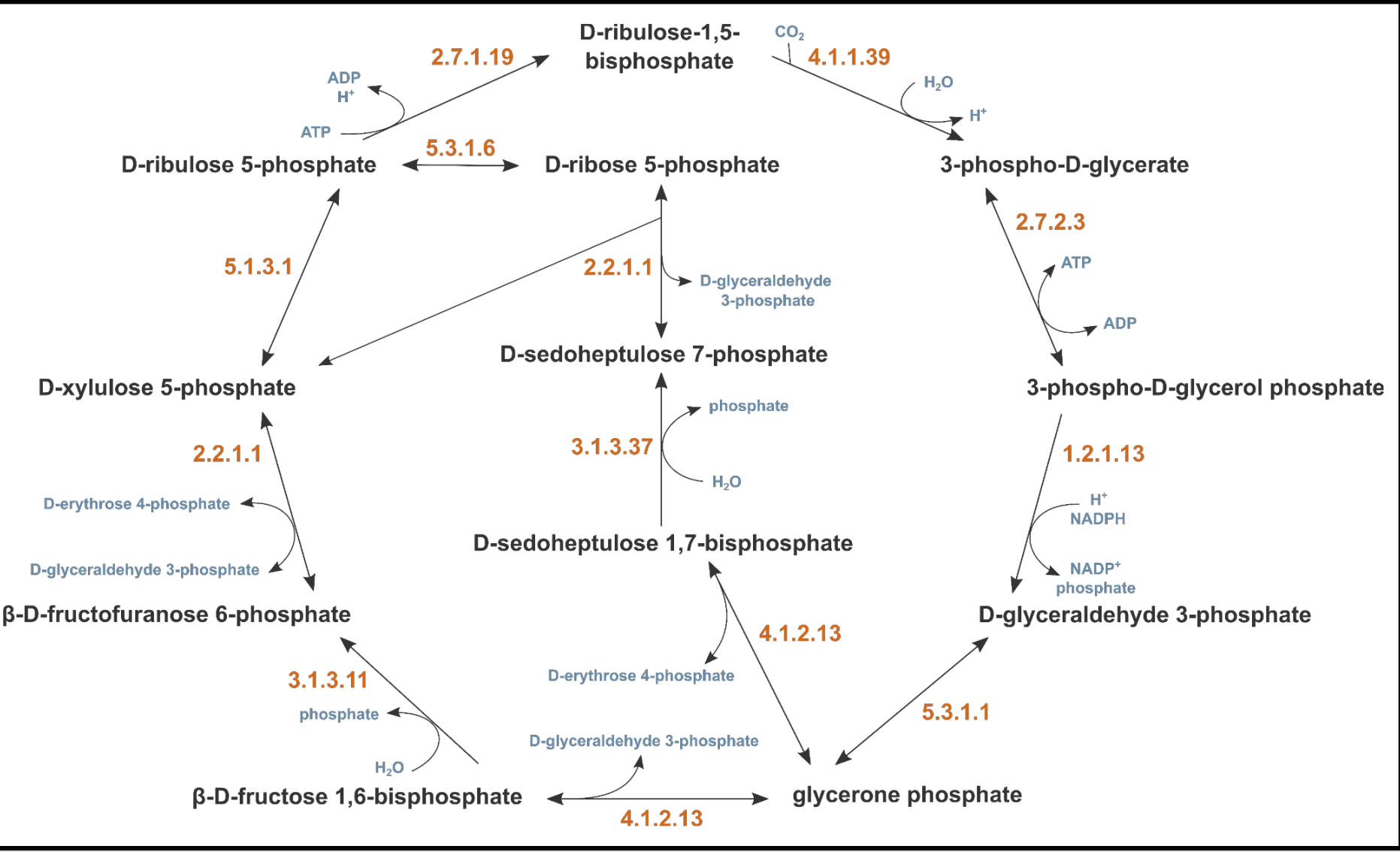

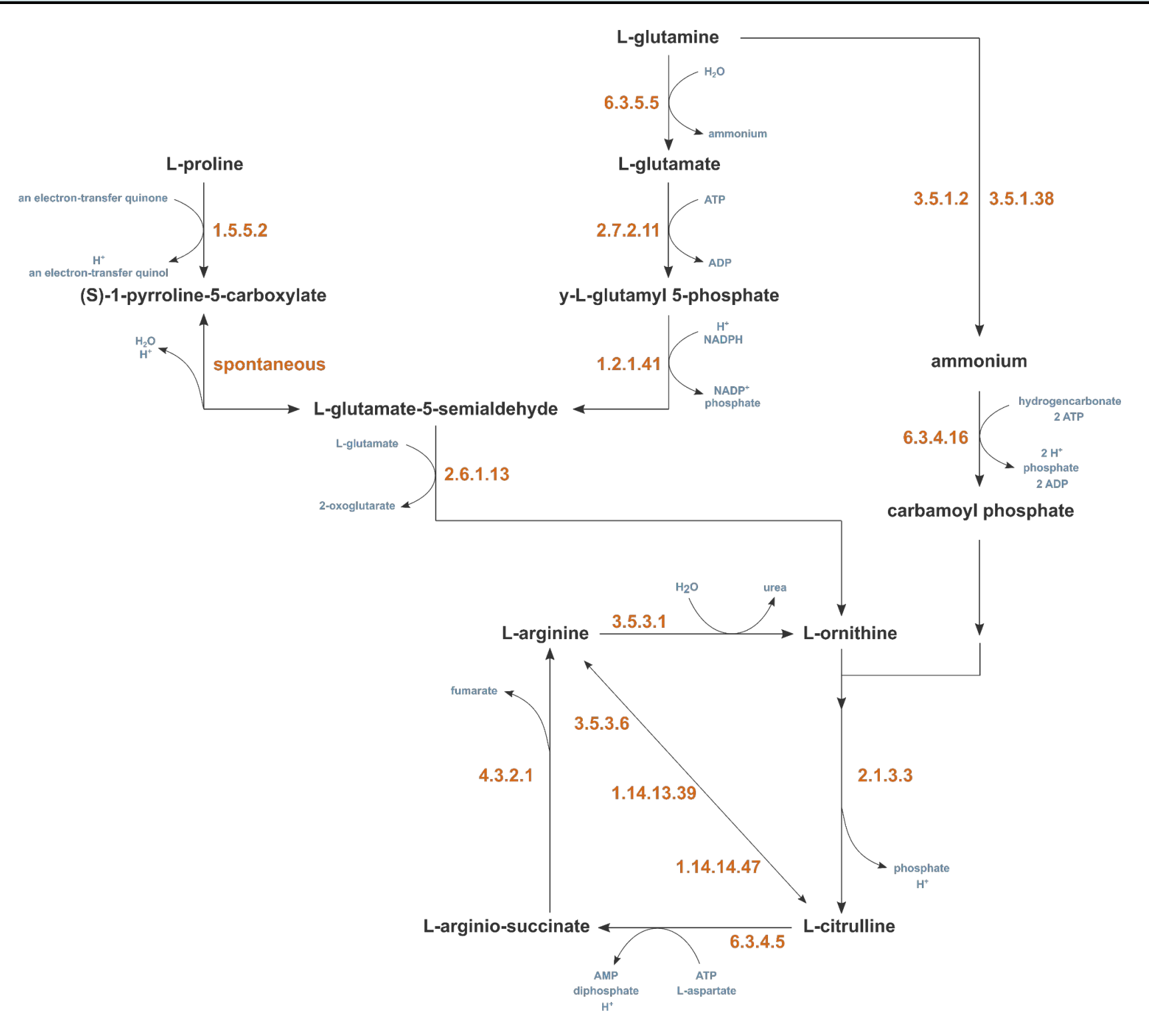Fig 1A: Calvin-Benson-Bassham Cycle (CBB)

Steps for data visualization

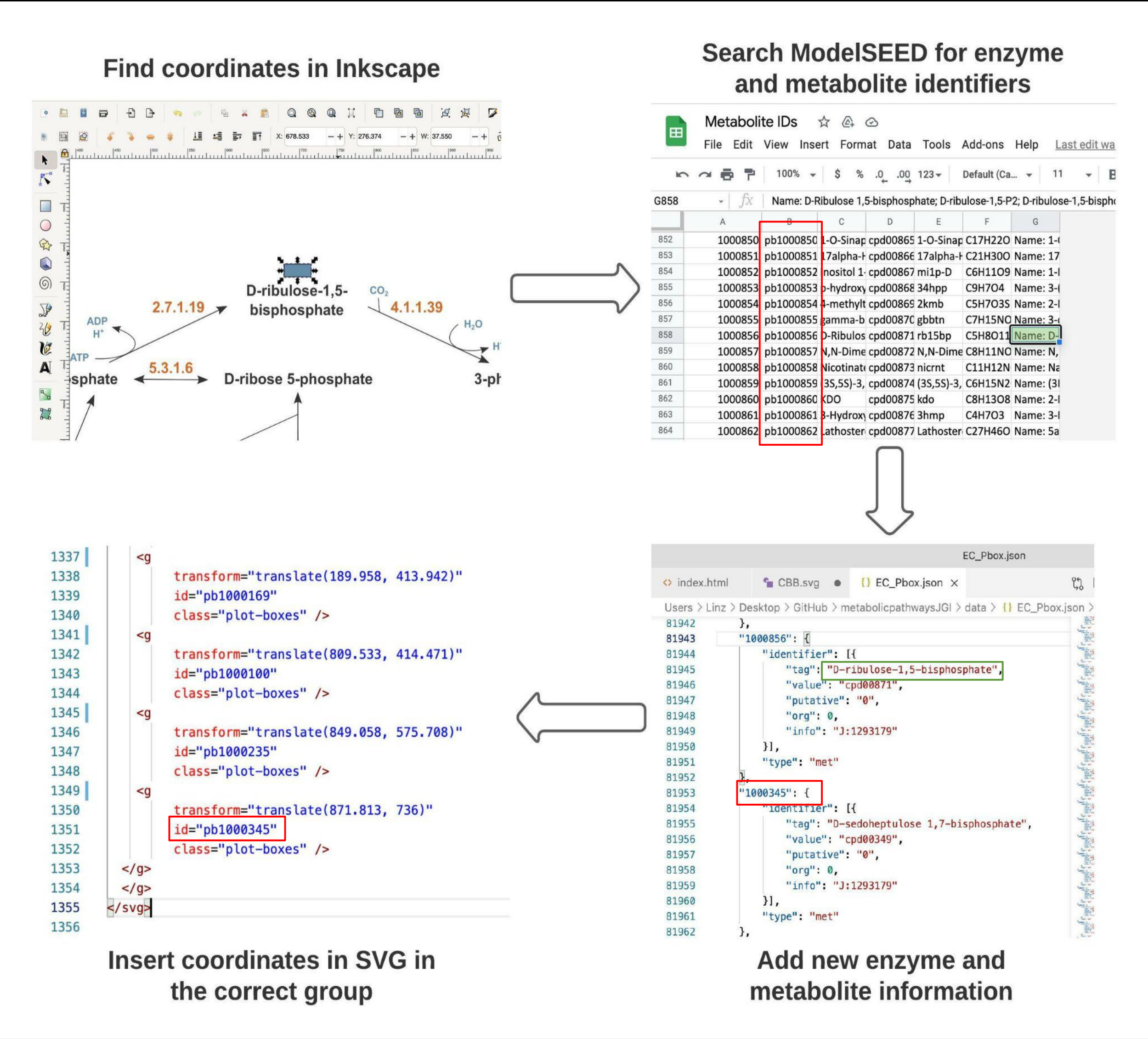Figure 2: Key steps for plotbox placement/data visualization using Inkscape, Visual Studio Code, and predefined databases containing metabolite and enzyme identifiers.

Fig 1B: Superpathway of L-Citrulline Metabolism
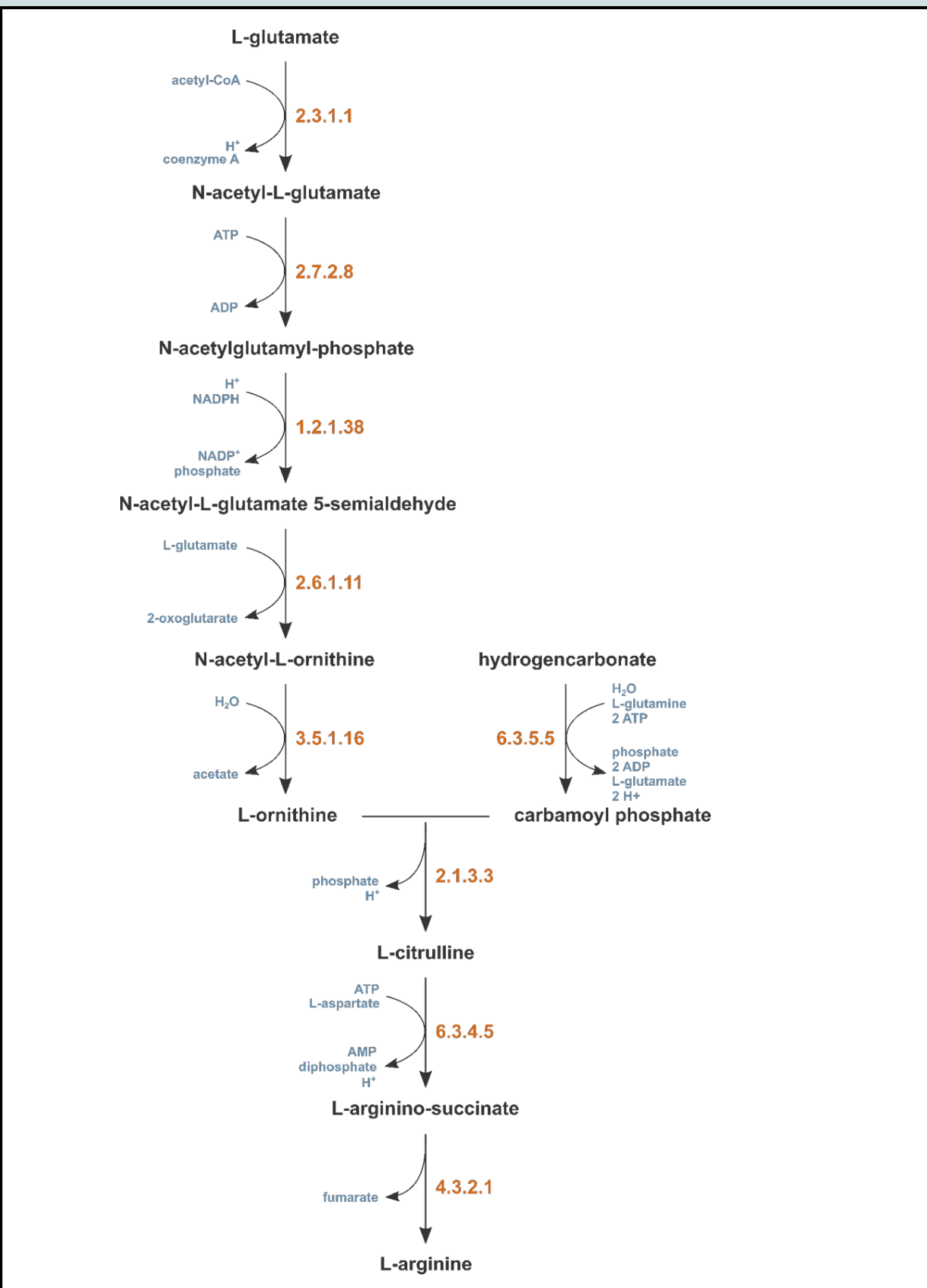
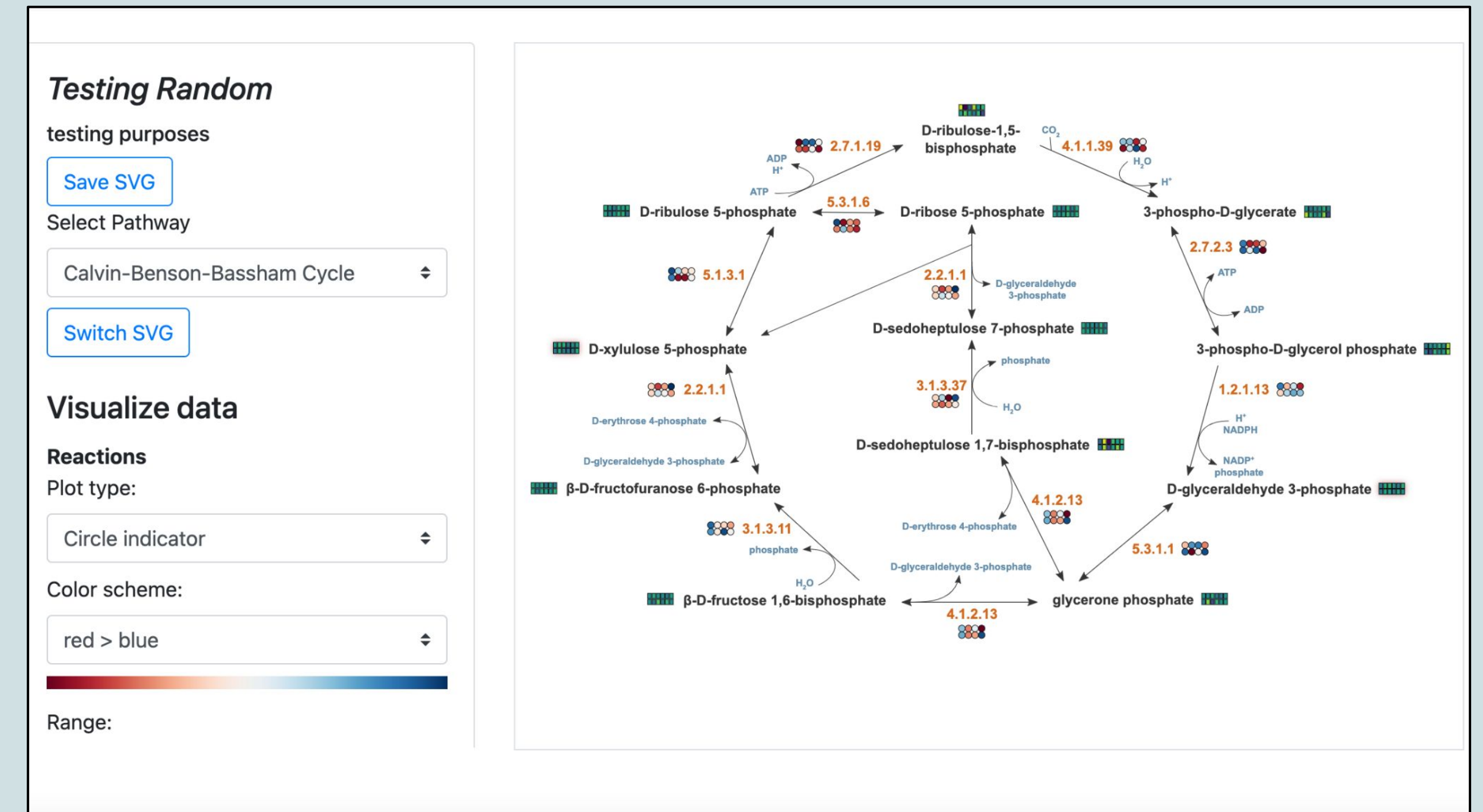Fig 1C: L-Arginine Biosynthesis I (via L-ornithine) pathway

Figure 3: CBB cycle (see Fig 1A) in MetaboMaps with multi-omics data projected. Transcriptomic data are represented by the circle indicators next to the enzymes and metabolomic data are represented by heatmaps next to the metabolites.

Background: Data visualization combined with multi-omics datasets allows scientists to check hypotheses on activation/regulation of metabolism. The ability to map multidimensional data onto simple visual representations of metabolic pathways allows researchers to grapple with such datasets without the need to utilize complex computational tools. Vector graphics can be used to represent network relations between enzymes and metabolites. When integrated into a web-based data analysis and visualization platform, they can be used to overlay multi-omics data on the pathway maps.

Results: Over thirty pathway maps were developed with embedded code to allow multi-omics data to be projected onto them. Final figures after rendering in the developed web app show both metabolomic and transcriptomic data (see Fig 3). These maps were imported to MetaboMaps. Some pathways were more complex than others, but the procedures to streamline the drawing process were gradually developed throughout.

JGI JOINT GENOME INSTITUTE